

US008245016B2

(12) **United States Patent**  
**Bybell**

(10) **Patent No.:** **US 8,245,016 B2**  
(45) **Date of Patent:** **\*Aug. 14, 2012**

(54) **MULTI-THREADED PROCESSING**

(75) Inventor: **Anthony J. Bybell**, Carrboro, NC (US)

(73) Assignee: **International Business Machines Corporation**, Armonk, NY (US)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 713 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **11/863,452**

(22) Filed: **Sep. 28, 2007**

(65) **Prior Publication Data**

US 2009/0089553 A1 Apr. 2, 2009

(51) **Int. Cl.**  
**G06F 7/00** (2006.01)

(52) **U.S. Cl.** ..... **712/217**

(58) **Field of Classification Search** ..... None  
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

7,584,342 B1 \* 9/2009 Nordquist et al. .... 712/22  
2004/0215892 A1 \* 10/2004 Fluhr et al. .... 711/137  
2004/0268093 A1 \* 12/2004 Samra et al. .... 712/217  
\* cited by examiner

*Primary Examiner* — Eddie P Chan

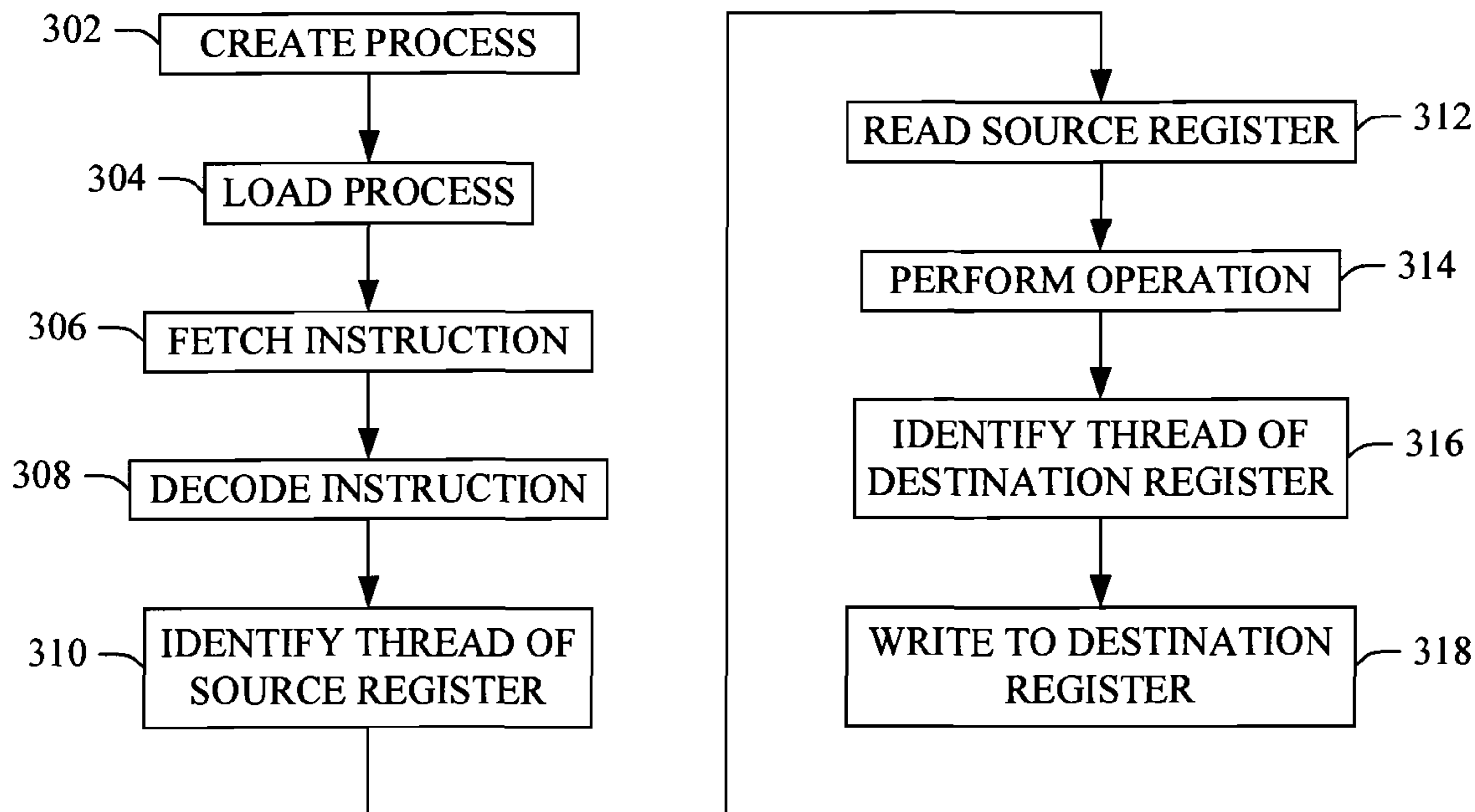
*Assistant Examiner* — John Lindlof

(74) *Attorney, Agent, or Firm* — Patrick J. Daugherty; Driggs, Hogg, Daugherty & Del Zoppo Co., LPA

(57) **ABSTRACT**

A system includes a multi-threaded processor that executes an instruction of a process of an executing program. The multi-threaded processor includes at least a first and a second thread. First and second sets of source registers are respectively allocated to the first and second threads, and first and second sets of destination registers are respectively allocated to the first and second threads. A resource prefix configuration register includes mappings between each of the source and destination registers and the threads. The multi-threaded processor, during execution of the instruction by one of the first or the second threads of execution, accesses the source and destination registers based on the mapping, wherein at least one of the accessed registers is allocated to the other of the first or the second thread of execution.

**18 Claims, 2 Drawing Sheets**



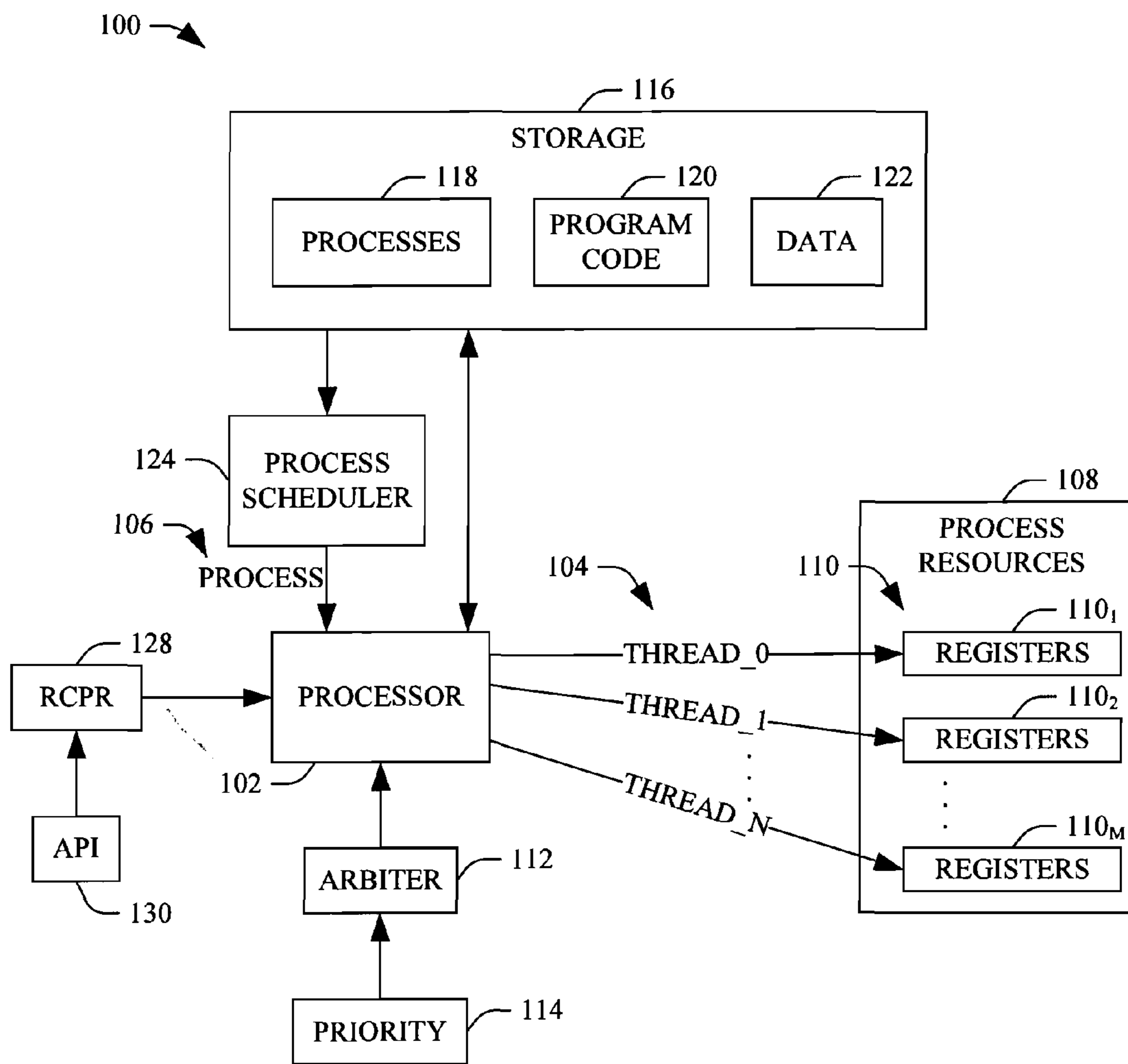


FIGURE 1

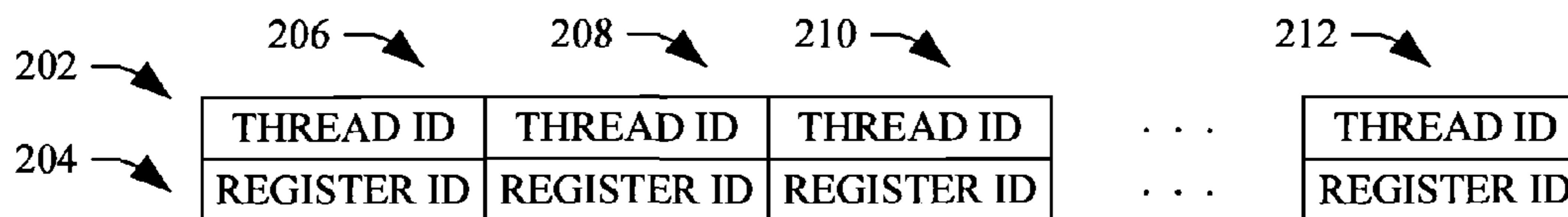


FIGURE 2

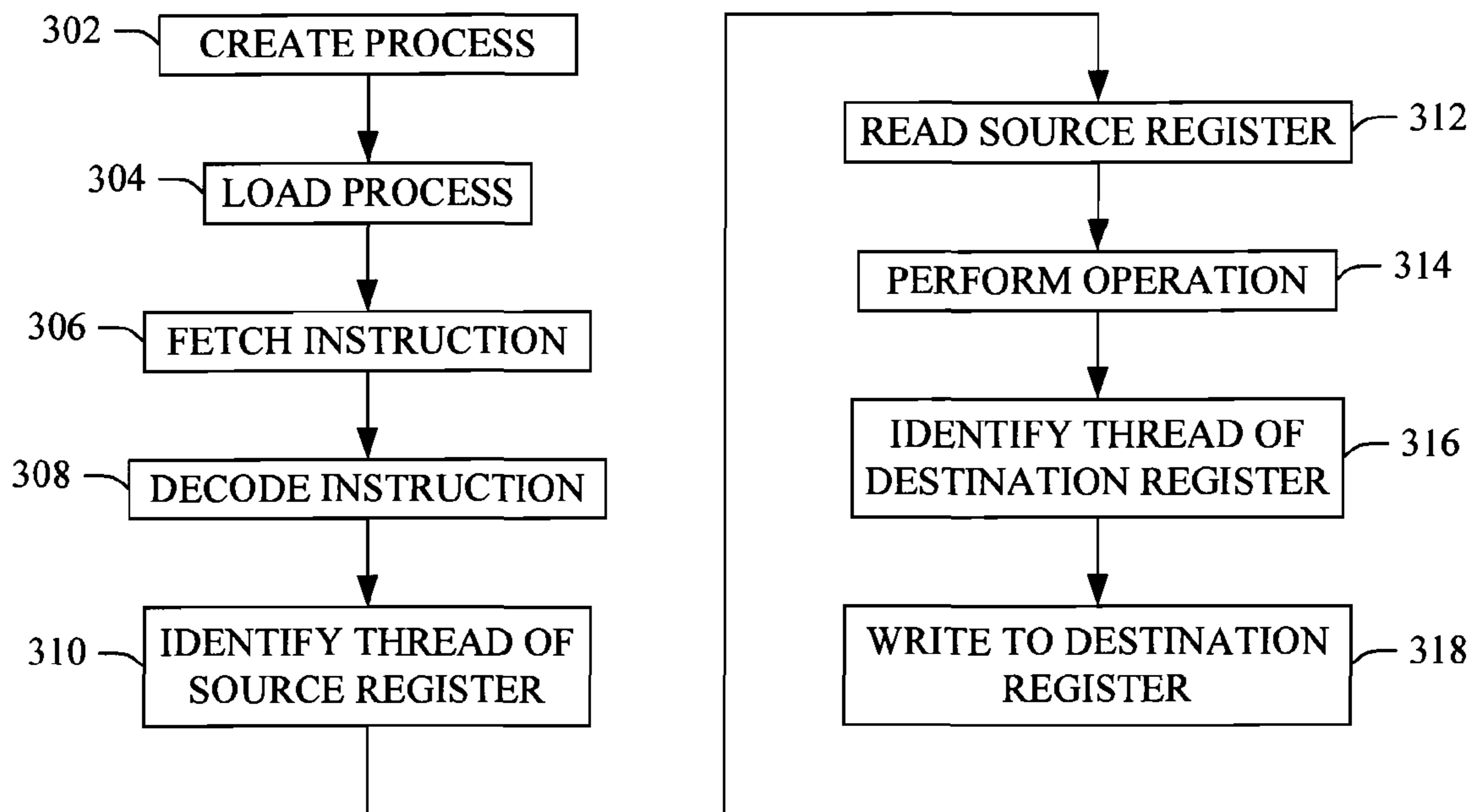


FIGURE 3

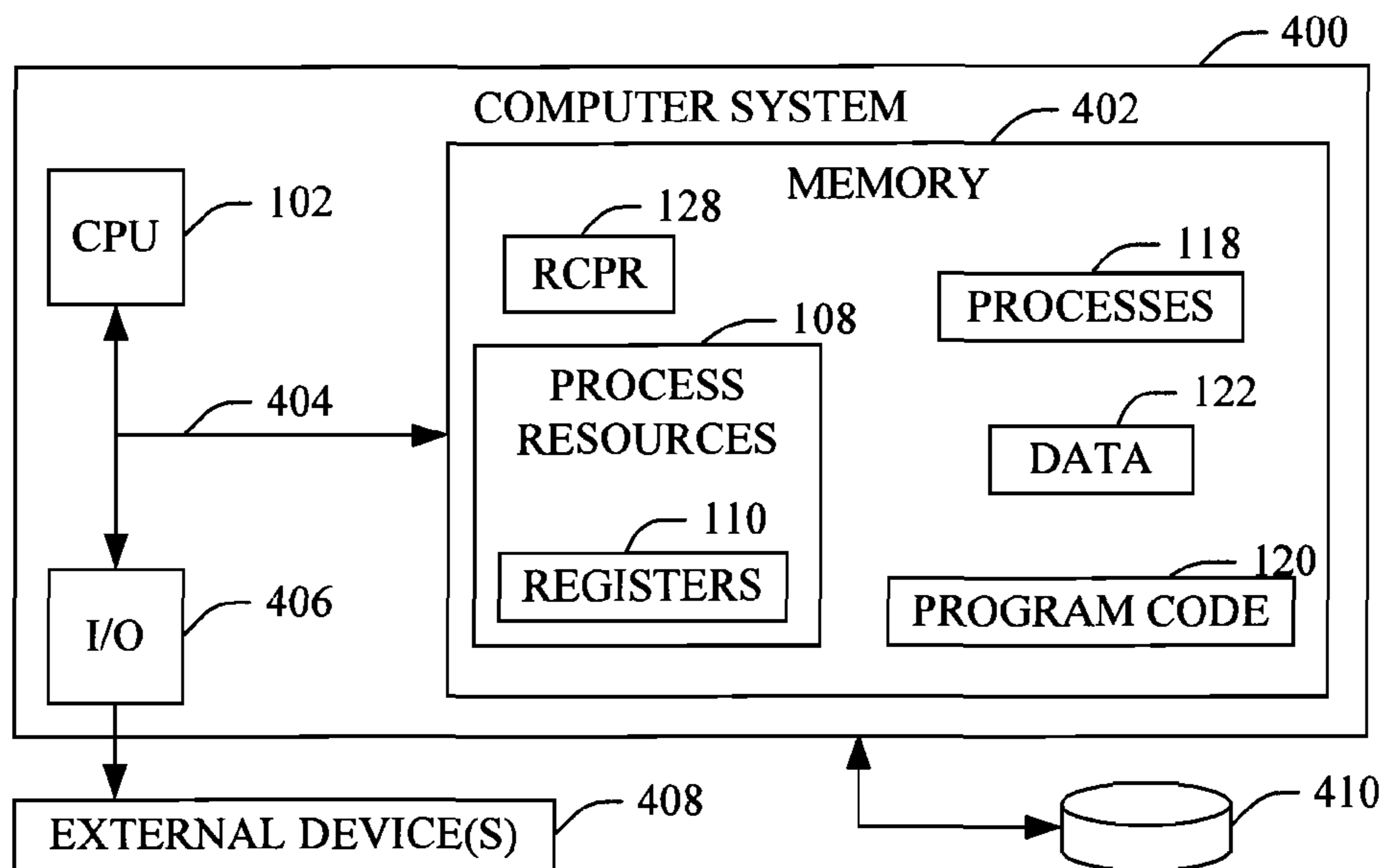


FIGURE 4

**MULTI-THREADED PROCESSING**

## FIELD OF THE INVENTION

The following generally relates to multithreaded processing and, more particularly, to a system and method for using thread resources of an inactive thread, in addition to the thread resources of an active thread, when executing an instruction with the active thread of the multithreaded processor.

## BACKGROUND OF THE INVENTION

With a multi-threaded processor, multiple threads of execution exist within the context of each process. The threads of a particular process are executed in a manner in which the processor quickly switches between different threads such that it appears that threads are being simultaneously executed. A simple type of multi-threading is where one thread runs until an event, such as a cache-miss that has to access off-chip memory, which might create a long latency. Rather than waiting, the processor switches to another thread that is ready to run. When the data for the previous thread arrives, the previous thread is placed back on the list of ready-to-run threads. In another type of multi-threading, the processor switches threads every CPU cycle.

Each process is allocated resources such as registers by the operating system, and such resources are allocated to the process' threads such that each thread owns its own resources, which are used when a thread is employed to execute an instruction. When a process is created, it is stored in main memory. Once the kernel assigns the process to a processor, the process is loaded into the processor and the processor executes the process's instructions using its resources.

A thread arbiter and/or thread priority determines which thread of execution to use to execute an instruction, and a thread identifier (ID) is associated with and follows the instruction through its various states of execution. The instruction is executed using the resources, such as the registers, of the thread that corresponds to the thread ID. When switching threads, the thread arbiter or thread priority determines the next thread to employ, and a thread ID of the next thread is associated with and follows the next instruction through its various states of execution. Likewise, the instruction is executed using the resources of the thread that corresponds to the thread ID of the next thread.

## SUMMARY OF THE INVENTION

In one aspect, a system includes a multi-threaded processor that executes an instruction of a process of an executing program. The multi-threaded processor includes at least a first thread of execution and a second thread of execution. The system further includes a plurality of source and destination registers, wherein a first set of the source registers is allocated to the first thread of execution, a second set of the source registers is allocated to the second thread of execution, a first set of the destination registers is allocated to the first thread of execution, and a second set of the destination registers is allocated to the second thread of execution. The system further includes a resource prefix configuration register that includes mapping between each of the source and destination registers and the threads of execution. The multi-threaded processor, during execution of the instruction by one of the first or the second threads of execution, accesses the source and destination registers based on the mapping, wherein at least one of the accessed registers is allocated to the other of the first or the second thread of execution.

In another aspect, a multi-threading method includes using a first of a plurality of threads of a processor to execute an instruction of a process, providing a first mapping between a source register identified in the instruction and a second of the plurality of threads, reading an operand from the register identified in the instruction in the second of the plurality of threads, and performing an operation identified in the instruction on the operand.

In another aspect, a multi-threaded processor includes at least two threads of execution, wherein the multi-threaded processor uses one of the at least two threads of execution to execute an instruction of a process, and the one of the at least two threads of execution utilizes resources allocated to the one of the at least two threads of execution and utilizes resources allocated to the other of the at least two threads of execution to execute the instruction.

## BRIEF DESCRIPTION OF THE DRAWINGS

These and other features will be more readily understood from the following detailed description of various examples taken in conjunction with the accompanying drawings in which:

FIG. 1 illustrates a system that includes a multi-threaded processor;

FIG. 2 illustrates a mapping between source and destination thread registers and a particular thread of execution for an instruction;

FIG. 3 illustrates a multi-threading method; and

FIG. 4 illustrates an example of a computer system in which the multi-threading method can be employed.

The drawings are merely representations and are not intended to portray specific elements. The drawings are intended for explanatory purposes and should not be considered as limiting scope.

## DETAILED DESCRIPTION OF THE INVENTION

FIG. 1 illustrates part of a computing system **100** such as a computer or other system that includes one or more processors, which may reside on the same or different dies. For sake of brevity and clarity, the illustrated system **100** is shown with a single, multithreaded processor **102**. The processor **102** can perform various functions, such as, but not limited to, executing instructions, exchanging data with the system's memory and/or peripherals, managing the system's other components, etc.

The illustrated processor **102** includes N threads of execution **104**, thread\_0, thread\_1, . . . , thread\_N. The processor **102** may use one or more of the threads of execution **104** when executing an instruction. For instance, the processor **102** may be run in single thread mode in which only one of the N threads is active. In another instance, the processor **102** may be run in multi-thread mode in which two to N threads are active. As discussed in greater detail below, active threads can use their respective resources as well as the resources of inactive threads when executing an instruction. In one instance, this allows access to the resources of the threads **104** in a non-orthogonal fashion and may increase the number of registers available to an active thread.

For each process, the operating system allocates resources **108**, which include, but are not limited to, various registers, file handles, security attributes, processor state, etc. Such resources **108** are shared across the threads **104** of a process. As shown, the resources **108** include multiple registers **110**, wherein a first set of registers **110<sub>1</sub>** has been allocated to the

## 3

thread\_0 a second set of registers **110**<sub>2</sub> has been allocated to the thread\_1, and an Mth set of registers **110**<sub>M</sub> has been allocated to the thread\_N.

As will be appreciated, the registers **110** may include one or more of general purpose registers (GPR's), floating point registers (FPR's), special purpose registers (SPR), etc. In addition, the number of the registers **110** and the number of bytes in each of the registers **110** may vary. For example, in one instance, at least one of the sets of registers **110**<sub>1</sub>, **110**<sub>2</sub>, . . . , **110**<sub>M</sub> includes thirty-two, 64-bit registers. Moreover, the registers **110** may be physical registers or copies of registers.

A storage component **116**, which may include or be part of main memory, stores information such as one or more created processes **118** for an executing program, program code **120**, data **122**, as well as other computer readable information.

A process scheduler **124** schedules the one or more processes **118** for execution by the processor **102**. As shown in this example, the process scheduler **124** has scheduled a process **106** for execution and, as such, the process **106** is loaded into the processor **102**. As noted above, when multiple processors are available the kernel first determines which processor will handle the process and then the process is loaded to the respective processor.

A resource prefix configuration register (RPCR) **128**, which may be a single register as shown or a set of registers, provides a mapping between the register(s) in an instruction and one or more of the threads of execution **104**. Such a mapping is used to prepend a thread ID to the resource specified by a given opcode field in an instruction. Using such a mapping, a thread can use its resources and/or the resources of one or more inactive threads when executing the instruction. As such, the total number of registers available to an active thread may be increased.

It is to be appreciated that this mapping may override the single thread ID that is typically generated in a system that does not employ the RPCR **128**. For example, a thread arbiter **112** may be used to determine which thread **104** is used by the processor **102** when the processor **102** executes an instruction. As shown, the thread arbiter **112** may use a thread priority **114** to facilitate identifying the thread **104** to be employed by the processor **102** for a particular instruction. The system can be configured so that the mapping provided by the RPCR **128** selectively overrides the arbiter **112**.

An application programming interface (API) **130** provides an interface that allows a user to define the mapping in the RPCR **128**. It is to be appreciated that the compiler may alternatively be used to define the mapping.

In operation, once the kernel has identified the processor **102** and the process **106** is loaded in the processor **102**, the processor **102** executes the process' instructions. This includes fetching an instruction of the process and decoding the instruction. Generally, such an instruction may include the operation to be performed and, if any, the source register of the operand(s) to be operated on and/or the destination register for the result. The operation may include, for example, an arithmetic operation, data copying operation, a logical operation, a program control operation, etc. The processor **102** then executes the decoded instruction.

As can be appreciated, in some instances executing the instruction may include reading data from one or more registers, operating on the data, and/or writing the results of the operation to one or more registers. In such instances, the processor **102** utilizes the mapping from the RPCR **128** to determine which thread and which thread register to read from and/or write to.

## 4

FIG. 2 illustrates a non-limiting example of such a mapping. In FIG. 2, a row **204** includes a plurality of fields for identifying a register via a register identifier (REGISTER ID) and a row **202** includes a plurality of fields for identifying a thread via a thread identifier (THREAD ID). Each column **206**, **208**, **210**, . . . , **212** maps a REGISTER ID to a THREAD ID.

By way of example, the process **106** may include the following instruction: XOR RA, RS, RB, wherein XOR represents the exclusive OR logical operation, RA represents the destination or result general purpose register, and RS and RB represent the source or operand general purpose registers. Execution of this instruction includes retrieving the operands from the RS and RB registers, XORing them together, and storing the result in the RA register. With reference to the example mapping in FIG. 2 and the above example instruction, a possible mapping for the instruction is depicted in Table 1 below.

TABLE 1

Example RCPR mapping.					
Thread ID:	00	00	01	00	...
Register ID:	RT	RS	RA	RB	...

With this mapping, the general purpose source registers RS and RB are mapped to thread\_0 (Thread ID 00) and the general purpose destination register RA is mapped to thread\_1 (Thread ID 01). As such, the instruction will result in reading the operands from the general purpose source registers RS and RB of thread\_0 (Thread ID 00), performing the XOR operation with the operands, and writing the result of the XOR operation to the general purpose destination register RA of thread\_1 (Thread ID 01). It is to be appreciated that the above is only an example. In another implementation the source registers RS and RB may be from different threads, one of the source registers RS or RB may be from the same thread as the destination register RA, etc. Of course, other instructions are also contemplated herein.

In the above example, the general purpose registers RS, RA, and RB are utilized. With other instructions, other general purpose registers such as RT, floating point registers such as FRT, FRA, FRB, FRC, etc., and/or special purpose registers such as SPR may additionally or alternatively be used.

When operating in single thread mode, the processor **102** executes the instruction via one thread (e.g., thread\_0), which reads and/or writes to the registers of one or more of the threads **104**, including its registers and/or one or more of the registers of one or more of the inactive threads thread\_1, thread\_N. In this mode, the processor **102** is provided with a copy of and/or access to the mapping in the RPCR **128**. As noted above, in one example each thread has thirty-two (32) 64-bit registers. By overloading the thread ID using the mapping, the active thread of a two-threaded processor may use up to sixty-four (64) registers, the active thread of a four-threaded processor may use up to one hundred and twenty-eight (128) registers, . . . .

In multi-thread mode, each active thread includes a copy of and/or has access to the mapping in the RPCR **128**. For example, assume that the processor **102** includes four (N=4) threads of operation, thread\_0, thread\_1, thread\_2, and thread\_3. Also assume that only two (2) of the threads, thread\_0 and thread\_3, are active, and that thread\_0 uses the resources of thread\_0, thread\_1 and thread\_2, and that thread\_3 only uses the resources of thread\_3. Both of the active threads, thread\_0 and thread\_3 would be provided with a copy of and/or access to the mapping in the RPCR **128**. Such information can be used by the software and/or hardware to

## 5

mitigate attempts by thread\_0 to access the resources used by thread\_3 (the resources of thread\_3), and attempts by thread\_3 to access the resources used by thread\_0 (the resources of thread\_0, thread\_1 and thread\_2).

In another example, active thread\_0 only uses the resources of inactive thread\_2, and active thread\_1 only uses the resources of inactive thread\_3. In this instance, both of the active threads, thread\_0 and thread\_1 would be provided with a copy of and/or access to the mapping in the RCPR 128. Such mapping may be used when one or more of the resources of thread\_0 and thread\_1 are corrupt. Various other permutations are also contemplated herein.

FIG. 3 illustrates a multi-threading method. The following is described in the context of a multi-threaded processor executing in single thread mode and an instruction that includes an operation, at least one operand source register, and at least one destination register. However, in other examples, more than one processor may be used, one or more of such processors may alternatively be ran in multi-threaded mode, and the instruction may alternatively only include a data copying operation, a program control operation, etc.

At reference numeral 302, a process is created, for example, for an executing program. At 304, the process is loaded in the processor. As noted above, with a multiprocessor system, the kernel first selects the processor that will execute the process. At 306, the processor fetches an instruction of the process. The schedule processor 124 or other technique may be used to determine process ordering. At 308, the processor decodes the instruction, including determining any operations, source registers and/or destination registers. At 310, the processor utilizes a mapping between the source registers and the threads of execution to identify which of the threads to read from to obtain an operand(s).

At 312, the processor reads the operand(s) from the source register(s) of the identified thread(s). At 314, the processor performs the operation on the operand(s). At 316, the processor utilizes a mapping between the destination registers and the threads of execution to identify which thread's register to write the result. At 318, the processor writes the result in the identified register of the identified thread. The processor fetches the next instruction, for example, as determined by the process scheduler 124, if any, and repeats 306 through 318. Otherwise, a next process, if any, is loaded in the processor.

Referring now to FIG. 4, an exemplary computer system 400 in which the system 100 can be utilized with is illustrated. As shown, the computer system 400 includes the multi-threaded processor (CPU) 102, memory 402, a bus 404, and input/output (I/O) interfaces 406. The memory 402 includes the RCPR 128 or a copy of the RCPR 128, the processes 118, the data 122, the program code 120, and the process resources 108, including the registers 110.

Further, the computer system 400 is shown in communication with external devices 408 and an external storage system 410. The external devices 408 can comprise any devices (e.g., keyboard, pointing device, display, etc.) that enable a user to interact with computer system 400 and/or any devices (e.g., network card, modem, etc.) that enable computer system 400 to communicate with one or more other computing devices. The external storage system 410 includes memory such as a hard disk, a floppy disk, portable memory, etc.

In general, the CPU 102 executes instructions of processes. While executing such instructions, the CPU 102 can read and/or write data to/from the memory 402, such as to/from the registers 110 of the process resources 108. As discussed above, different sets of the registers 110 allocated to a process are allocated to each of the threads of the CPU 102. The CPU 102 accesses the RCPR 128 or the copy of the RCPR 128 to

## 6

obtain a mapping between the registers identified in the instruction and the different threads of execution. The mapping provides an override that allows a thread of execution executing an instruction to use the resources 108, including the registers 110, of inactive threads of execution, along with its own resources, to facilitate execution of the instruction.

The foregoing description of various aspects of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and obviously, many modifications and variations are possible. Such modifications and variations that may be apparent to a person skilled in the art are intended to be included within the scope of the invention as defined by the accompanying claims.

I claim:

1. A multi-threading processing method, the method comprising:

mapping respective identifying fields of each of a plurality of source registers and of a plurality of destination registers to different respective identifying fields of each of a plurality of threads of instruction; wherein a first of the source registers and a first of the destination registers are mapped to a first thread; wherein a second of the source registers and a second of the destination registers are mapped to a second thread; and wherein the first thread is different from the second thread, and each of the first and second source registers and the first and second destination registers are different from each other of the first and second source registers and the first and second destination registers; and

executing an instruction of the first thread via a processor when the second thread is inactive by:

providing access by the first thread to the mapping of the second thread to the second source register and the second destination register if the second thread is inactive; using the mapping of the second thread to the second source register and the second destination register to execute the instruction of the first thread by reading from the second source register and writing to the second destination register;

using the mapping of the first thread to the first source register and the first destination register to execute the instruction of the first thread by reading from the first source register and writing to the first destination register; and

mitigating any attempt by the second thread to read from the second source register or write to the second destination register still mapped to the second thread while the first thread is using said second source register to execute the instruction of the first thread.

2. The method of claim 1, further comprising:

storing the mapping in a resource prefix configuration register in communication with the processor; and

wherein the using the mapping to read from the second source register and write to the second destination register comprises using the mapping stored in the resource prefix configuration register.

3. The method of claim 2, further comprising:

providing an interface that allows a user to define the mapping in the resource prefix configuration register.

4. The method of claim 2, further comprising:

using a compiler to define the mapping in the resource prefix configuration register.

5. The method of claim 1, further comprising:

using a thread priority to identify via a thread arbiter the first thread executed by the processor; and

7

selectively overriding via the thread arbiter the mapping of the second source register and the second destination register to the second thread and mitigating the any attempt by the second thread to read from the second source register or write to the second destination register still mapped to the second thread while the first thread is using said second source register and said second thread is inactive.

6. The method of claim 1, wherein the pluralities of source and destination registers include one or more of a general purpose register, a floating point register, and a special purpose register.

7. A multi-threading system, the system comprising: a multithreaded processor that maps via respective identifying fields of each of a plurality of source registers and of a plurality of destination registers to different respective identifying fields of each of a plurality of threads of instruction; wherein a first of the source registers and a first of the destination registers are mapped to a first thread; wherein a second of the source registers, and a second of the destination registers are mapped to a second thread; and wherein the first thread is different from the second thread, and each of the first and second source registers and the first and second destination registers are different from each other of the first and second source registers and the first and second destination registers; and

wherein the processor executes an instruction of the first thread when the second thread is inactive by:

providing access by the first thread to the mapping of the second thread to the second source register and the second destination register if the second thread is inactive; using the mapping of the second thread to the second source register and the second destination register to execute the instruction of the first thread by reading from the second source register and writing to the second destination register;

using the mapping of the first thread to the first source register and the first destination register to execute the instruction of the first thread by reading from the first source register and writing to the first destination register; and

mitigating any attempt by the second thread to read from the second source register or write to the second destination register still mapped to the second thread while the first thread is using said second source register to execute the instruction of the first thread.

8. The system of claim 7, further comprising: a resource prefix configuration register in communication with the processor, wherein the resource prefix configuration register stores the mapping; wherein the processor uses the mapping stored in the resource prefix configuration register to read from the second source register and write to the second destination register.

9. The system of claim 8, further comprising: an application programming interface in communication with the resource prefix configuration register, wherein the application programming interface provides an interface that allows a user to define the mapping in the resource prefix configuration register.

10. The system of claim 8, further comprising: a compiler in communication with the resource prefix configuration register, wherein the compiler defines the mapping in the resource prefix configuration register.

8

11. The system of claim 7, further comprising:

a thread arbiter in communication with the processor, wherein the thread arbiter determines which thread is used by the processor when the processor executes an instruction; and

wherein the thread arbiter uses a thread priority to facilitate identifying the first thread executed by the processor; and

wherein the thread arbiter selectively overrides the mapping during the executing of the instruction of the first thread via the processor when the second thread is inactive.

12. The system of claim 7, wherein the pluralities of source and destination registers include one or more of a general purpose register, a floating point register, and a special purpose register.

13. A computer program product for multi-threaded processing, the computer program product comprising:

a non-transitory computer readable storage medium;

first program instructions to map respective identifying fields of each of a plurality of source registers and of a plurality of destination registers to different respective identifying fields of each of a plurality of threads of instruction; wherein a first of the source registers and a first of the destination registers are mapped to a first thread; wherein a second of the source registers, and a second of the destination registers are mapped to a second thread; and wherein the first thread is different from the second thread, and each of the first and second source registers and the first and second destination registers are different from each other of the first and second source registers and the first and second destination registers; and

second program instructions to execute an instruction of the first thread when the second thread is inactive by:

providing access by the first thread to the mapping of the second thread to the second source register and the second destination register if the second thread is inactive; using the mapping of the second thread to the second source register and the second destination register to execute the instruction of the first thread by reading from the second source register and writing to the second destination register;

using the mapping of the first thread to the first source register and the first destination register to execute the instruction of the first thread by reading from the first source register and writing to the first destination register; and

mitigating any attempt by the second thread to read from the second source register or write to the second destination register still mapped to the second thread while the first thread is using said second source register to execute the instruction of the first thread; and

wherein the first and second program instructions are stored on the computer readable storage medium.

14. The computer program product of claim 13, wherein the first program instructions are further to store the mapping in a resource prefix configuration register in communication with the processor and to use the mapping stored in the resource prefix configuration register to read from the second source register and write to the second destination register.

15. The computer program product of claim 14, wherein the first program instructions are further to provide an interface that allows a user to define the mapping in the resource prefix configuration register.

9

16. The computer program product of claim 14, wherein the first program instructions are further to use a compiler to define the mapping in the resource prefix configuration register.

17. The computer program product of claim 13, wherein the first program instructions are further to provide a thread arbiter that determines which thread is used by the processor when the processor executes an instruction and uses a thread priority to identify the first thread executed by the processor; and

wherein the second program instructions are further to selectively override the mapping of the second source register and the second destination register to the second

10

thread during the executing of the instruction of the first thread when the second thread is inactive and to mitigate the any attempt by the second thread to read from the second source register or write to the second destination register still mapped to the second thread while the first thread is using said second source register and said second thread is inactive.

18. The computer program product of claim 13, wherein the pluralities of source and destination registers include one or more of a general purpose register, a floating point register, and a special purpose register.

\* \* \* \* \*